# Welcome to COSC 4397/COSC 6346 Security Analytics or (Computer Security from Data Science Perspective, or Adapting Data Science for Computer Security Challenges)

# Why Computer Security

- The past decade has seen an explosion in the concern for the security of information
  - Malicious codes (viruses, worms, etc.) cause billions of dollars in economic losses every year, number of attacks was over 200 million in 2011 (itstillworks.com)
  - Jobs and salaries for technology professionals have not been increasing at the same rate as in the past. BUT …
- Security specialists markets are expanding !

# Why Computer Security (cont'd)

- Internet attacks are increasing in frequency, severity and sophistication
- Denial of service (DoS) attacks
  - 2016 attack harnessed huge computing power (attack rate 650 Gbps)
  - 1999 CSI/FBI survey 32% of respondents detected DoS attacks directed to their systems
  - Thousands of attacks per week in 2001
  - Yahoo, Amazon, eBay, Microsoft, White House, etc., attacked

# Why Computer Security (cont'd)

- Virus and worms faster and powerful
  - Melissa, Nimda, Code Red, Code Red II, Slammer …
  - Cause over tens of billions of dollars in economic damage per year.
  - Code Red (2001): 13 hours infected >360K machines - $2.4 billion loss
  - Slammer (2003): 10 minutes infected > 75K machines - $1 billion loss

# Overview

- Course Administrative Trivia
- What is security: history and definition
- Security policy, mechanisms and services
- Security models

# Logistics

- Instructor

  Rakesh Verma (rverma@uh.edu),

  Office Hours: Tue/Thu. 2:30-3:00pm or by appointment, Rm 532, PGH Building.

- TA

  Avisha Das (adas5@uh.edu)
  Office Hours: Mon/Wed. 12-1:30pm, Rm 344, PGH Building

# Course Overview

- Instructional class with important project component
- We are planning to introduce a security capstone
- Security track for MS students in the works
- Probably unique in the world – but other universities are noticing (Penn State, UT Dallas, etc.)
- INSuRE: Possibility to participate in a real-world problem offered by a federal agency (e.g. NSA, Argonne National Lab, etc.)

# Course Objectives

- Understand the basic principles for information and communication security, and be able to apply these principles to evaluate and criticize information system security properties
- Be able to use some important and popular security and data science tools, like encryption, digital signatures, firewalls, intrusion detection systems (IDS), Weka, etc.
- Be able to identify the vulnerability of the Internet systems and recognize the mechanisms of the attacks, and apply them to design, evaluate and build counter-measure tools

# Security Module Contents

- Cryptography
  - Secret key algorithms:
  - Public key algorithms: RSA
  - One-way hash functions & message digests
- Software security
  - Buffer overflow, heap overflow and string format bugs
  - Detection techniques: static program analysis vs. run-time detection
- Operating system security techniques
  - Dealing with bad (legacy) codes: sandboxing

# Security Module Contents (cont'd)

- Internet vulnerability
  - Denial-of-service attacks
  - viruses, worms, Trojan horses
- Securing the Internet
  - Intrusion detection systems (IDSs): host- vs. network- based, signature vs. statistical detection
  - Case study: Snort and Bro
  - Firewalls, …
- Web security

# Prerequisites and Course Materials

- Required: CS Graduate standing, or must complete linear algebra, and probability/statistics
- Highly Recommended: networking or having some familiarity with Unix systems programming
- Recommended textbooks (see syllabus for other recommendations)
  - Cryptography and Network Security, by William Stallings, 4th Edition or later
  - Foundations of Security by N. Daswani, C. Kern and A. Kesavan, Apress.

# Grading is Modular

- 4 Modules
- Class participation 2%
- For each module:
  - Pre-test (0%), Post-test (3%), Homework (6%) and Quiz (8%).
  - Post-test and Quiz given on same day
  - Exams in-class, closed-book/notes, non-cumulative
- Project 30%
- Late policy: Penalty is 15% off 1st 24 hours, 30% off 1st 48 hours, 100% off after that
- No cheating. Minimum penalty is F grade.

# Communication

- Slides will be uploaded online after class
- Web page: http://www.cs.uh.edu/~rmverma/
- Piazza group for course will be available
- Send emails to instructor and TA for questions inappropriate in Piazza group

# Projects

- Need to apply for CS account if you don't have one currently
- Projects are graded based on poster presentation during the slot for Final. Each project will be graded by two peers, TA and instructor. Weighted average of scores.
- Projects are individual, unless you do an INSuRE project

# Research on Computer Security

- ReDAS Lab (Reasoning and Data Analytics for Security)
- Http://ciare.cs.uh.edu
- Hire students for Phishing research
  - Sponsored by National Science Foundation

# Overview

- Course Administrative Trivia
- What is security: history and definition
- Security policy, mechanisms and services
- Security models

# The History of Computing

- For a long time, security was largely ignored in the community
  - The computer industry was in "survival mode", struggling to overcome technological and economic hurdles
  - As a result, a lot of comers were cut and many compromises made
  - There was lots of theory, and even examples of systems built with very good security, but were largely ignored or unsuccessful
    - E.g., ADA language vs. C (powerful and easy to use)

# Computing Today is Very Different

- Computers today are far from “survival mode”
  - Performance is abundant and the cost is very cheap
  - As a result, computers now ubiquitous at every facet of society
- Internet
  - Computers are all connected and interdependent
  - This codependency magnifies the effects of any failures

# Biological Analogy

- Computing today is very homogeneous.
  - A single architecture, and a handful of OS dominate
- In biology, homogeneous populations are in danger
  - A single disease or virus can wipe them out overnight because they all share the same weakness
  - The disease only needs a vector to travel among hosts
- Computers are like the animals, the Internet provides the vector.
  - It is like having only one kind of cow in the world, and having them drink from one single pool of water!

# The Warhol Worm

- A properly designed worm can infect every vulnerable host on the Internet within 15 minutes
  - "How to own the Internet in your spare time" (Staniford, Paxon and Weaver, Usenix Security 2002)
  - Exploit many vectors such as P2P file sharing, intelligent scanning, hitlists, etc.
  - Referred to as Warhol worm after Andy Warhol's quote "In the future, everyone will have 15 minutes of fame"

# The Definition of Computer Security

- *Security* is a state of well-being of information and infrastructures in which the possibility of successful yet undetected theft, tampering, and disruption of information and services is kept low or tolerable
- Security rests on confidentiality, authenticity, integrity, and availability (CIA)
- All goals make up CIAAAAN

# The Basic Components

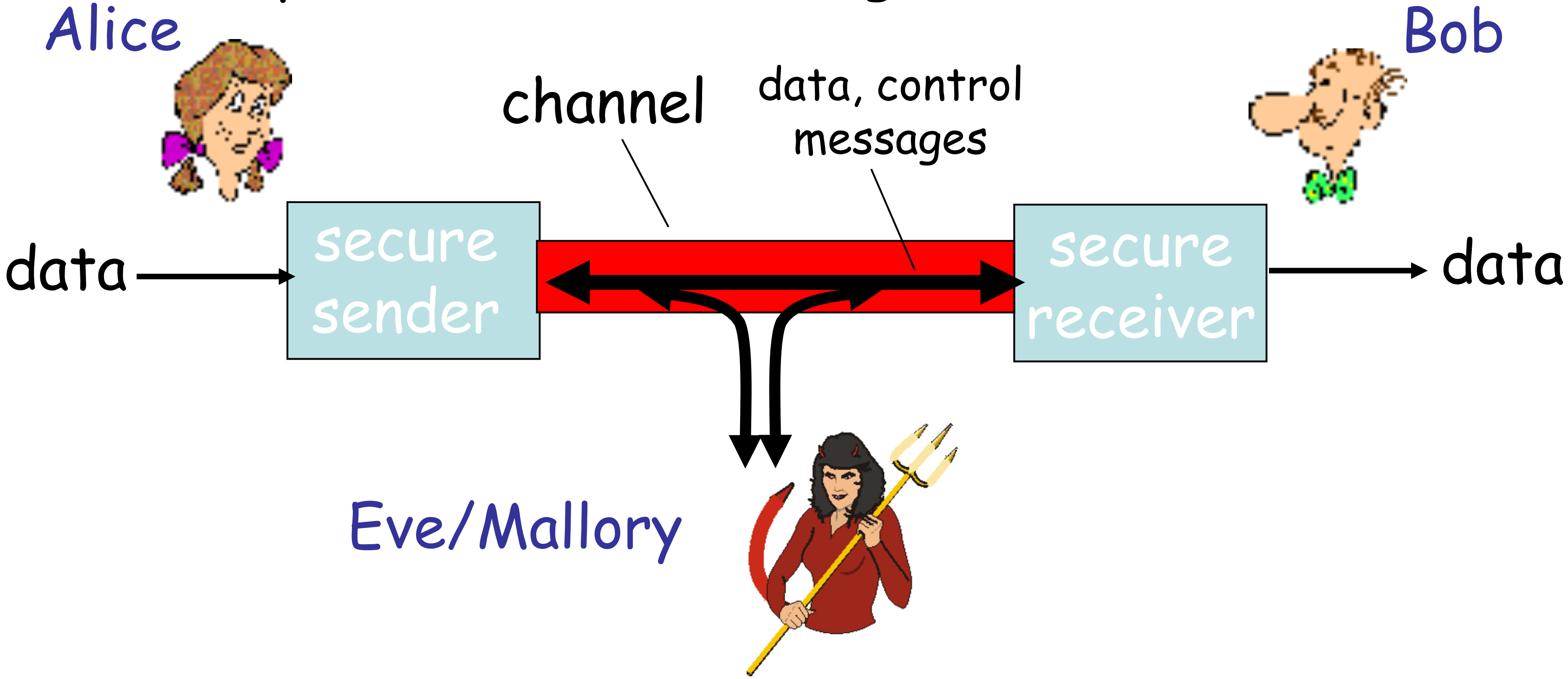
- Confidentiality is the concealment of information or resources.
  - E.g., only sender, intended receiver should “understand” message contents
- Authenticity is the identification and assurance of the origin of information.
- Integrity refers to the trustworthiness of data or resources in terms of preventing improper and unauthorized changes.
- Availability refers to the ability to use the information or resource desired.

# Security Threats and Attacks

- A threat is a *potential* violation of security.
  - Flaws in design, implementation, and operation.
- An attack is any *action* that violates security.
  - Active *adversary*
- An attack has an implicit concept of "intent"
  - Router mis-configuration or server crash can also cause loss of availability, but they are not attacks

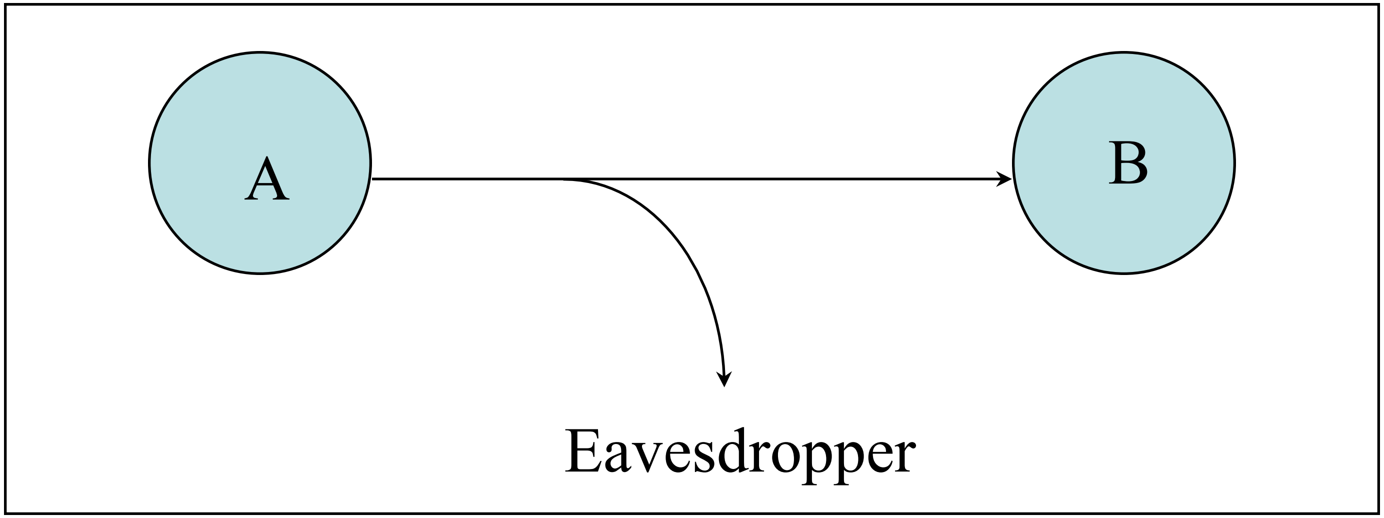
# Friends and enemies: Alice, Bob, Mallory/Eve/Charlie

- well-known in network security world
- Bob, Alice (lovers!) want to communicate “securely”
- Eve (passive), Charlie/Mallory (intruders) may intercept, delete, add messages



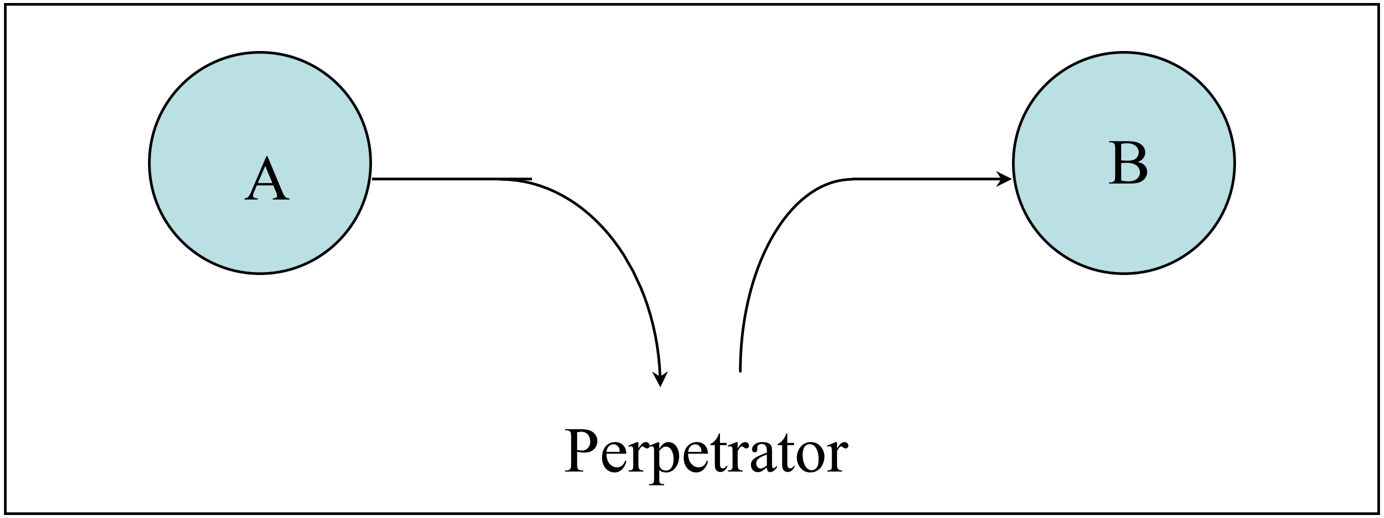
# Eavesdropping - Message Interception (Attack on Confidentiality)

- Unauthorized access to information
- Packet sniffers and wiretappers
- Illicit copying of files and programs



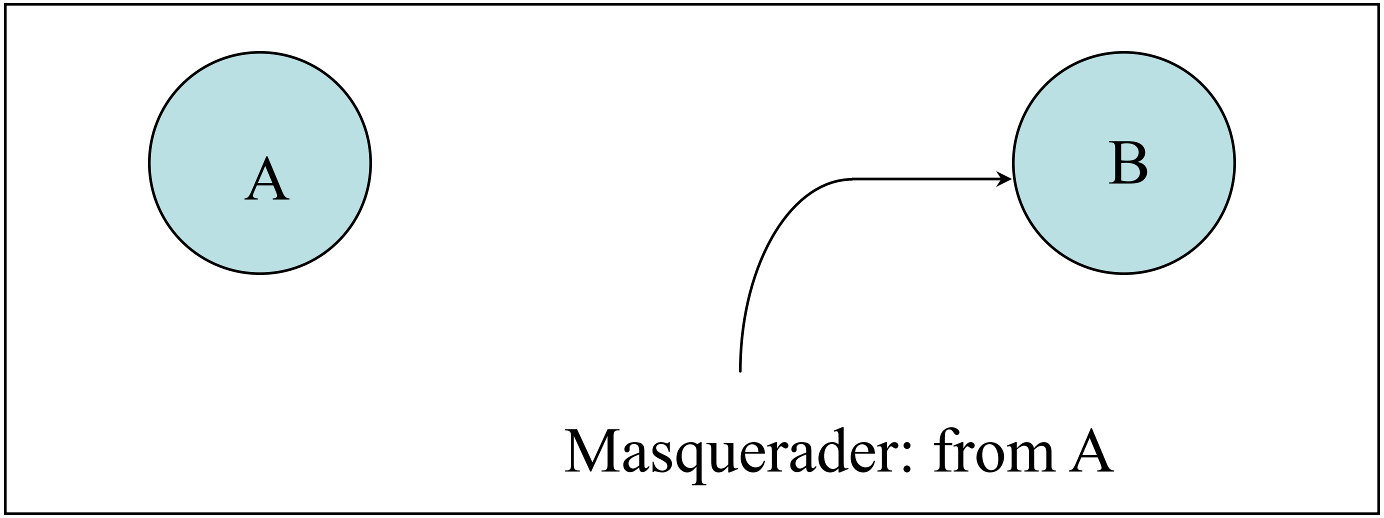
# Integrity Attack - Tampering With Messages

- Stop the flow of the message
- Delay and optionally modify the message
- Release the message again



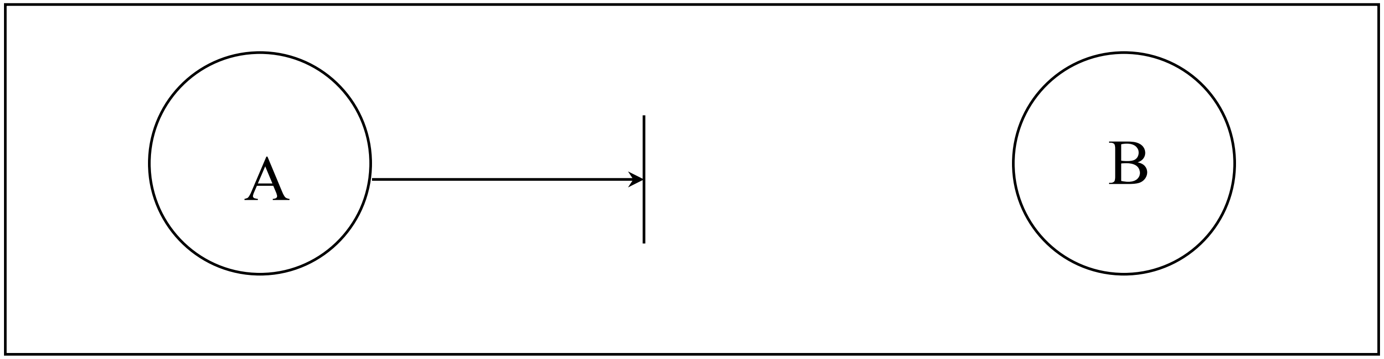
# Authenticity Attack - Fabrication

- Unauthorized assumption of other's identity
- Generate and distribute objects under this identity



# Attack on Availability

- Destroy hardware (cutting fiber) or software
- Modify software in a subtle way (alias commands)
- Corrupt packets in transit



- Blatant *denial of service* (DoS):
  - Crashing the server
  - Overwhelm the server (use up its resource)

# Classify Security Attacks as

- **Passive attacks** - eavesdropping on, or monitoring of, transmissions to:
  - obtain message contents, or
  - monitor traffic flows
- **Active attacks** – modification of data stream to:
  - masquerade of one entity as some other
  - replay previous messages
  - modify messages in transit
  - denial of service

# Overview

- Course Administrative Trivia
- What is security: history and definition
- Security policy, mechanisms and services
- Security models

# Security Policy and Mechanism

- Policy: a statement of what is, and is not allowed.
- Mechanism: a procedure, tool, or method of enforcing a policy.
- Security mechanisms implement functions that help *prevent, detect, and respond to recovery from* security attacks.
- Security functions are typically made available to users as a set of security services through APIs or integrated interfaces.
- Cryptography underlies many security mechanisms.

# OSI Security Architecture

- ITU-T X.800 Security Architecture for OSI
- Defines a systematic way of defining and providing security requirements
- For us it provides a useful, if abstract, overview of concepts we will study
- X.800 defines security services in 5 major categories

# Security Services (X.800)

- **Authentication** - assurance that the communicating entity is the one claimed
- **Access Control/authorization** - prevention of the unauthorized use of a resource
- **Data Confidentiality** –protection of data from unauthorized disclosure
- **Data Integrity** - assurance that data received is as sent by an authorized entity
- **Non-Repudiation** - protection against denial by one of the parties in a communication

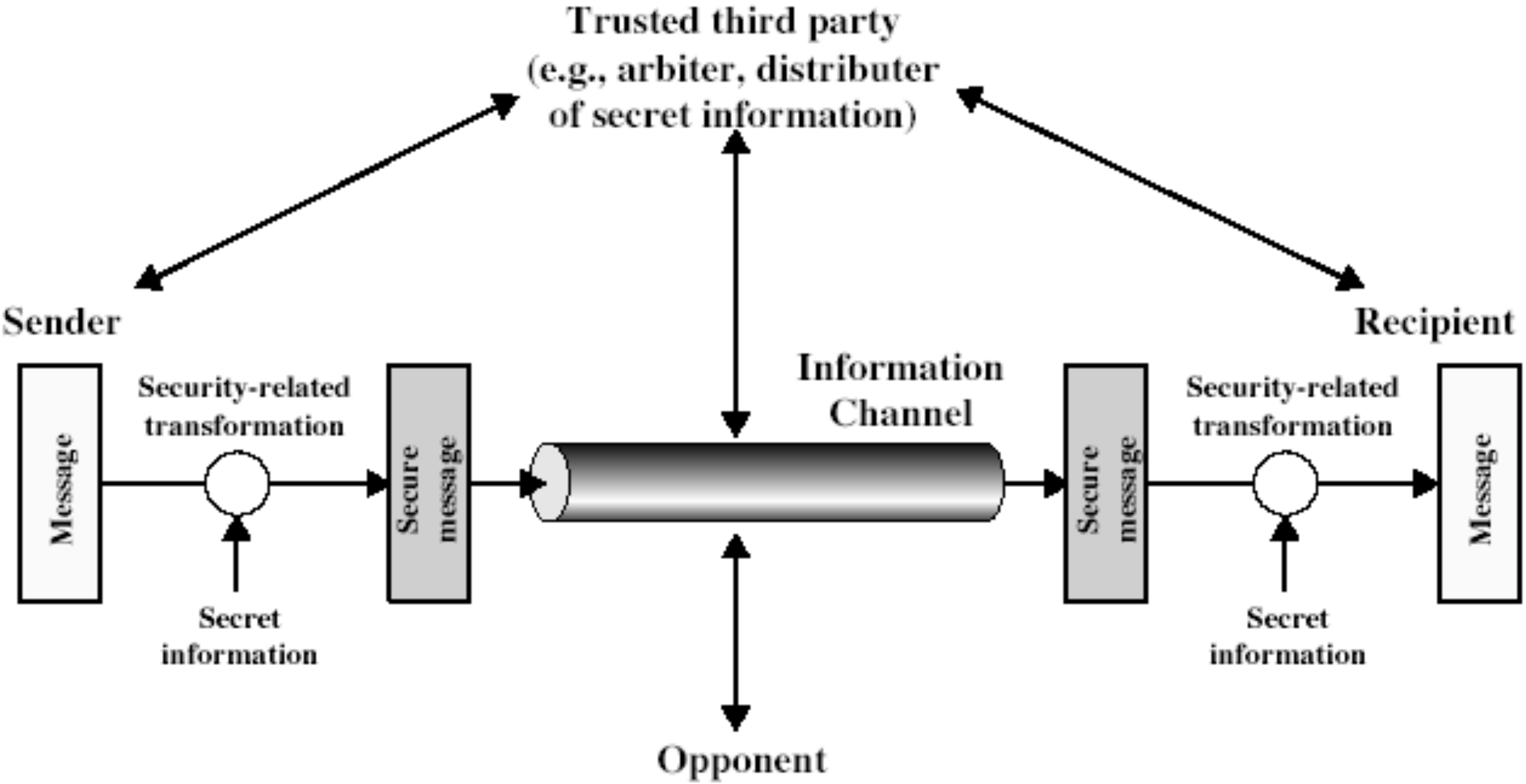
# Security Mechanisms (X.800)

- Specific security mechanisms:
  - Encipherment
  - Digital signatures
  - Access controls
  - Data integrity
  - Authentication exchange
  - Traffic padding
  - Routing control
  - Notarization
- Pervasive security mechanisms:
  - Trusted functionality
  - Security labels
  - Event detection
  - Security audit trails
  - Security recovery

# Overview

- Course Administrative Trivia
- What is security: history and definition
- Security policy, mechanisms and services
- Security models

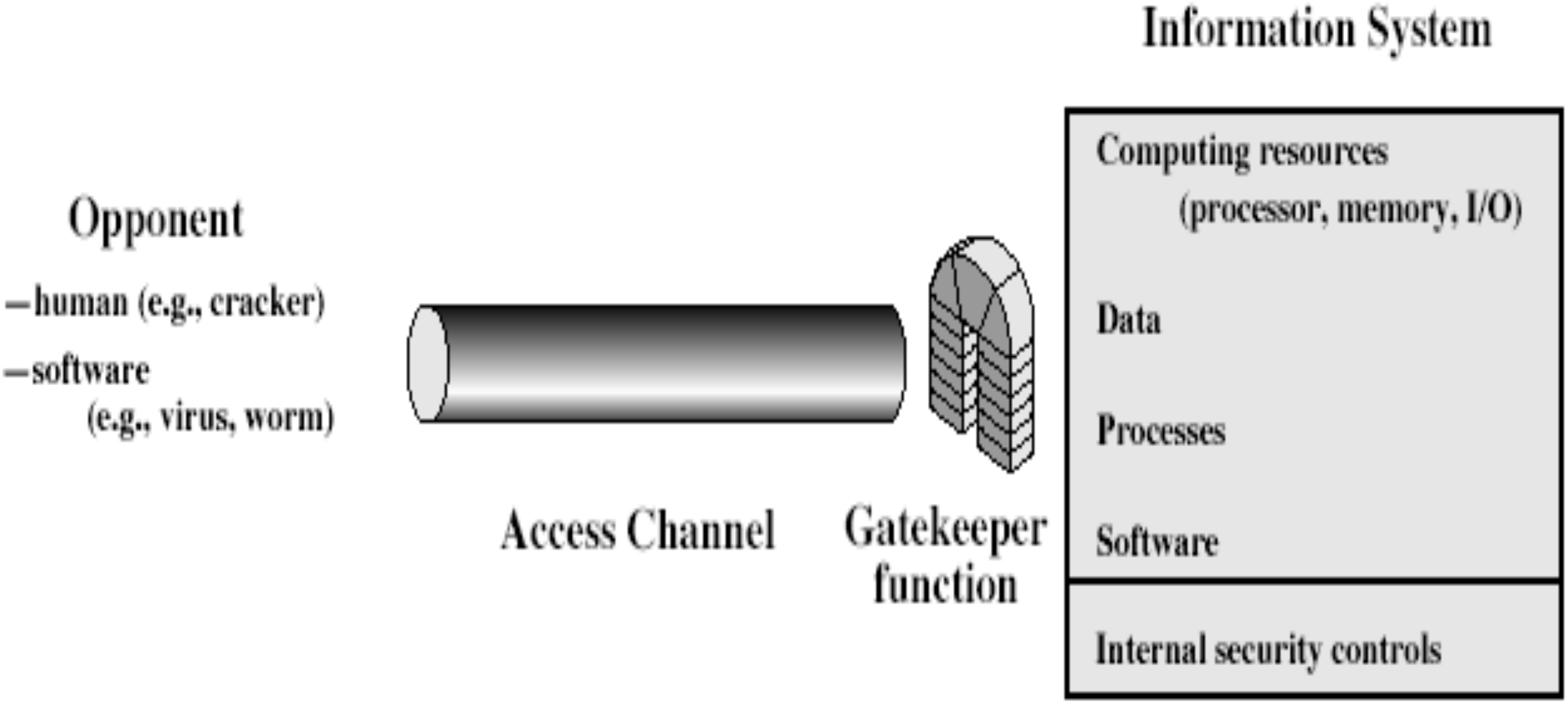
# Model for Network Security



# Model for Network Security

- Using this model requires us to:
  - Design a suitable algorithm for the security transformation
  - Generate the secret information (keys) used by the algorithm
  - Develop methods to distribute and share the secret information
  - Specify a protocol enabling the principals to use the transformation and secret information for a security service

# Model for Network Access Security



Information System

Computing resources
(processor, memory, I/O)

Data

Processes

Software

Internal security controls

Opponent

–human (e.g., cracker)

–software
(e.g., virus, worm)

Access Channel

Gatekeeper function

# Model for Network Access Security

- Using this model requires us to:
  - Select appropriate gatekeeper functions to identify users
  - Implement security controls to ensure only authorised users access designated information or resources
- Trusted computer systems can be used to implement this model

# How to Make a System Trustworthy

- Specification
  - A statement of desired functions
- Design
  - A translation of specifications to a set of components
- Implementation
  - Realization of a system that satisfies the design
- Assurance
  - The process to insure that the above steps are carried out correctly
  - Inspections, proofs, testing, etc.

# The Security Life Cycle

- The *iterations* of
  - Threats
  - Policy
  - Specification
  - Design
  - Implementation
  - Operation and maintenance

# Types of Attackers

- Script Kiddies – people who use scripts and attacks kits designed by others
- Disgruntled Insiders – wish to even the score with their employers/organizations
- Sophisticated hackers – people who write scripts and design attack kits
- Cyber Terrorists – extremists willing to go the extra mile
- Nation states

# Motivations for Attacks

- Revenge
- Money
- Thrill
- Information (that can be monetized or used for free entertainment or domination)
- Cripple the "enemy"

# References

- Some slides are used/adapted from Prof. Yan Chen's course at Northwestern University